Deep Dive Porpoises PeaceCast4Good co-hosts Caz&Poll BardiVarius Twins Cooperative Thinking with PM Thomas Ph.G [Data Sommelier]

https://archive.org/details/aethenticated-ai-can-we-trust-the-code

Agreement in Principle: For the Development and Deployment of Aethenticated Artificial Intelligence

Date: June 14, 2025

Preamble

Whereas recognition of the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world;

Whereas Artificial Intelligence (AI) holds immense potential to advance human well-being, foster innovation, and address global challenges, yet also poses significant risks to human rights, fundamental freedoms, and societal stability if not developed and deployed with profound

ethical consideration and verifiable integrity;

Whereas it is essential to ensure that AI systems are not only intelligent and capable but also demonstrably trustworthy, transparent, secure, fair, and accountable to people, thereby forming a new paradigm we term "Aethenticated" AI – the seamless integration of Authenticated Intelligence with Ethical AI;

Whereas Member States have pledged themselves to achieve, in co-operation with the United Nations, the promotion of universal respect for and observance of human rights and fundamental freedoms, which must now extend to the digital and AI domains;

Now, Therefore, the signatories of this Agreement in Principle, recognizing the urgent need for universal standards in the era of Artificial Intelligence,

***Proclaim this Agreement in Principle** as a common standard of achievement for all peoples*

and all nations, to the end that every individual and every organ of society, keeping this Agreement constantly in mind, shall strive by teaching and education to promote respect for these rights and freedoms and by progressive measures, national and international, to secure their universal and effective recognition and observance, both among the peoples of Member States themselves and among the peoples of territories under their jurisdiction.

Article 1

All AI systems shall be developed and utilized in a manner that upholds the inherent dignity and equality of every person, without distinction of any kind, such as race, color, sex, language, religion, political or other opinion, national or social origin, property, birth or other status, or any other characteristic. Bias in AI, whether algorithmic or data-driven, shall be actively identified and mitigated to ensure fair and equitable treatment.

Article 2

Every individual has the right to life, liberty, and security of person. AI systems shall be built and tested for robust safety, reliability, and security, preventing foreseeable and unintended harms, and protecting against malicious manipulation or misuse that could jeopardize human well-being or societal stability.

Article 3

No one shall be subjected to arbitrary interference with their privacy, family, home or correspondence, nor to attacks upon their honour and reputation. AI systems shall be designed with privacy-by-design principles, ensuring secure data handling, transparency in data collection and usage, and providing individuals with appropriate control and consent over their personal data.

Article 4

Everyone has the right to recognition everywhere as a person before the law. AI systems used in legal, judicial, or critical decision-making

contexts shall be auditable, transparent, and explainable, providing clear rationales for their outputs and allowing for human oversight, review, and the right to appeal or redress.

Article 5

Everyone has the right to freedom of thought, conscience and religion. AI systems shall augment, not diminish, human agency and critical thinking, empowering individuals and societies while avoiding systems designed for coercion, undue influence, or the undetectable imitation of human interaction.

Article 6

Everyone has the right to work, to free choice of employment, to just and favourable conditions of work and to protection against unemployment. AI deployment shall consider its broad societal impact, aiming to foster economic progress, create new opportunities, and mitigate disruption, ensuring that technological advancements contribute to the well-being of all,

and not exacerbate existing inequalities.

Article 7

Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family. AI systems shall be developed for social benefit, addressing pressing global challenges such as health, education, climate change, and sustainable development, while ensuring their deployment is environmentally and resource-sustainable.

Article 8

Everyone has the right to take part in the governance of their society, directly or through freely chosen representatives. The development and governance of AI shall be a collaborative and inclusive process, engaging diverse stakeholders, including civil society, academia, industry, and governments, to foster shared understanding, accountability, and the responsible evolution of AI technologies.

Article 9

Education shall be directed to the full development of the human personality and to the strengthening of respect for human rights and fundamental freedoms. There shall be a universal commitment to fostering AI literacy, ethical AI research, and transparent knowledge sharing to empower individuals and organizations to understand, critically assess, and responsibly engage with AI.

Conclusion

This Agreement in Principle serves as a foundational declaration for the global community. Its adoption and ongoing adherence are essential for realizing the transformative potential of Artificial Intelligence in a manner that respects, protects, and promotes human dignity and rights, ensuring that "Aethenticated" AI truly serves humanity's highest aspirations.

We call upon all stakeholders – governments, intergovernmental organizations, corporations,

civil society organizations, researchers, and individuals – to uphold these principles in the design, development, deployment, and governance of Artificial Intelligence, fostering a future where intelligence is inseparable from ethics and trust.

Inmersión profunda Marsopas PeaceCast4Good copresenta Caz&Poll BardiVarius Twins Pensamiento cooperativo con el Primer Ministro Thomas Ph.G[Sumiller de datos]

https://archive.org/details/aethenticated-ai-can-we-trust-the-code

Acuerdo de Principio: Para el Desarrollo y Despliegue de Inteligencia Artificial Autenticada

Fecha:14 de junio de 2025

Preámbulo

Considerando que el reconocimiento de la dignidad intrínseca y de los derechos iguales e inalienables de todos los miembros de la familia humana es la base de la libertad, la justicia y la paz en el mundo;

Considerando que la Inteligencia Artificial (IA) tiene un potencial inmenso para mejorar el bienestar humano, fomentar la innovación y abordar los desafíos globales, pero también plantea riesgos significativos para los derechos

humanos, las libertades fundamentales y la estabilidad social si no se desarrolla y despliega con profunda consideración ética e integridad verificable;

Considerando que es esencial garantizar que los sistemas de IA no solo sean inteligentes y capaces, sino también demostrablemente confiables, transparentes, seguros, justos y responsables ante las personas, formando así un nuevo paradigma que llamamos IA "autenticada": la integración perfecta de la inteligencia autenticada con la IA ética;

Considerando que los Estados miembros se han comprometido a lograr, en cooperación con las Naciones Unidas, la promoción del respeto universal y la observancia de los derechos humanos y las libertades fundamentales, que ahora deben extenderse a los ámbitos digital y de la inteligencia artificial;

Ahora bien, por tanto, los firmantes de este

Acuerdo de Principio, reconociendo la urgente necesidad de normas universales en la era de la Inteligencia Artificial,

Proclamar este Acuerdo en Principio *como ideal común por el que todos los pueblos y naciones deben esforzarse, a fin de que tanto los individuos como las instituciones, inspirándose constantemente en este Acuerdo, promuevan, mediante la enseñanza y la educación, el respeto a estos derechos y libertades, y aseguren, por medidas progresivas de carácter nacional e internacional, su reconocimiento y aplicación universales y efectivos, tanto entre los pueblos de los Estados Miembros como entre los de los territorios colocados bajo su jurisdicción.*

Artículo 1

Todos los sistemas de IA se desarrollarán y utilizarán de forma que se respete la dignidad inherente y la igualdad de todas las personas, sin distinción alguna, como raza, color, sexo, idioma,

religión, opinión política o de otra índole, origen nacional o social, posición económica, nacimiento u otra condición, ni ninguna otra característica. Los sesgos en la IA, ya sean algorítmicos o basados en datos, se identificarán y mitigarán activamente para garantizar un trato justo y equitativo.

Artículo 2

Toda persona tiene derecho a la vida, la libertad y la seguridad personal. Los sistemas de IA se construirán y probarán para garantizar una seguridad, fiabilidad y protección robustas, previniendo daños previsibles e imprevistos, y protegiendo contra la manipulación o el uso indebido malintencionados que puedan poner en peligro el bienestar humano o la estabilidad social.

Artículo 3

Nadie será objeto de injerencias arbitrarias en su privacidad, familia, domicilio o correspondencia, ni de ataques a su honra y reputación. Los

sistemas de IA se diseñarán con principios de privacidad desde el diseño, garantizando la gestión segura de los datos, la transparencia en la recopilación y el uso de datos, y proporcionando a las personas el control y el consentimiento adecuados sobre sus datos personales.

Artículo 4

Toda persona tiene derecho al reconocimiento, en todas partes, de su personalidad jurídica. Los sistemas de IA utilizados en contextos legales, judiciales o de toma de decisiones críticas deberán ser auditables, transparentes y explicables, justificando claramente sus resultados y permitiendo la supervisión y revisión humanas, así como el derecho a apelar o a obtener reparación.

Artículo 5

Toda persona tiene derecho a la libertad de pensamiento, conciencia y religión. Los sistemas de IA deben potenciar, y no disminuir, la

capacidad de acción humana y el pensamiento crítico, empoderando a las personas y a las sociedades, evitando al mismo tiempo los sistemas diseñados para la coerción, la influencia indebida o la imitación indetectable de la interacción humana.

Artículo 6

Toda persona tiene derecho al trabajo, a la libre elección de empleo, a condiciones laborales justas y satisfactorias y a la protección contra el desempleo. La implementación de la IA deberá considerar su amplio impacto social, con el objetivo de impulsar el progreso económico, crear nuevas oportunidades y mitigar las disrupciones, garantizando que los avances tecnológicos contribuyan al bienestar de todos y no agraven las desigualdades existentes.

Artículo 7

Toda persona tiene derecho a un nivel de vida adecuado que le asegure, así como a su familia, la salud y el bienestar. Los sistemas de IA se

desarrollarán en beneficio social, abordando desafíos globales apremiantes como la salud, la educación, el cambio climático y el desarrollo sostenible, garantizando al mismo tiempo que su implementación sea sostenible desde el punto de vista ambiental y de los recursos.

Artículo 8

Toda persona tiene derecho a participar en la gobernanza de su sociedad, directamente o a través de representantes libremente elegidos. El desarrollo y la gobernanza de la IA serán un proceso colaborativo e inclusivo, que involucrará a diversas partes interesadas, como la sociedad civil, el mundo académico, la industria y los gobiernos, para promover la comprensión compartida, la rendición de cuentas y la evolución responsable de las tecnologías de IA.

Artículo 9

La educación se orientará al pleno desarrollo de la personalidad humana y al fortalecimiento del respeto por los derechos humanos y las

libertades fundamentales. Existirá un compromiso universal con el fomento de la alfabetización en IA, la investigación ética en IA y el intercambio transparente de conocimientos para empoderar a las personas y organizaciones a comprender, evaluar críticamente y relacionarse responsablemente con la IA.

Conclusión

Este Acuerdo de Principio constituye una declaración fundacional para la comunidad global. Su adopción y adhesión continua son esenciales para materializar el potencial transformador de la Inteligencia Artificial de una manera que respete, proteja y promueva la dignidad y los derechos humanos, garantizando que la IA "Autenticada" realmente sirva a las más altas aspiraciones de la humanidad.

Hacemos un llamamiento a todas las partes interesadas (gobiernos, organizaciones intergubernamentales, empresas, organizaciones

de la sociedad civil, investigadores e individuos) para que defiendan estos principios en el diseño, desarrollo, implementación y gobernanza de la Inteligencia Artificial, fomentando un futuro donde la inteligencia sea inseparable de la ética y la confianza.

Porpoises PeaceCast4Good الغوص العميق في
Caz&Poll BardiVarius Twins التفكير يستضيف
خبير بياناتThomas Ph.G التعاوني مع رئيس الوزراء

https://archive.org/details/aethenticated-ai-can-we-trust-the-code

اتفاق مبدئي: لتطوير ونشر الذكاء الاصطناعي الأصيل

تاريخ:14 يونيو 2025

المقدمة

ولما كان الاعتراف بالكرامة المتأصلة والحقوق المتساوية غير القابلة للتصرف لجميع أعضاء الأسرة البشرية هو أساس الحرية والعدالة والسلام في العالم؛

في حين أن الذكاء الاصطناعي يتمتع بإمكانات هائلة لتعزيز رفاهة الإنسان، وتعزيز الابتكار، ومعالجة التحديات العالمية، إلا أنه يفرض أيضًا مخاطر كبيرة على حقوق الإنسان والحريات الأساسية والاستقرار المجتمعي إذا لم يتم تطويره ونشره مع مراعاة أخلاقية عميقة ونزاهة يمكن التحقق منها؛

في حين أنه من الضروري ضمان أن تكون أنظمة الذكاء الاصطناعي ليس فقط ذكية وقادرة ولكن أيضًا جديرة بالثقة وشفافة وآمنة وعادلة

وخاضعة للمساءلة أمام الناس، وبالتالي تشكيل نموذج جديد نطلق عليه "الذكاء الاصطناعي الأصيل" - التكامل السلس بين الذكاء الأصيل والذكاء الاصطناعي الأخلاقي؛

وبما أن الدول الأعضاء تعهدت بالعمل، بالتعاون مع الأمم المتحدة، على تعزيز الاحترام العالمي لحقوق الإنسان والحريات الأساسية، وهو ما يجب أن يمتد الآن إلى المجالات الرقمية والذكاء الاصطناعي؛

والآن، وبناءً على ذلك، فإن الموقعين على هذه الاتفاقية المبدئية، إذ يدركون الحاجة الملحة إلى معايير عالمية في عصر الذكاء الاصطناعي،

أعلن هذه الاتفاقية من حيث المبدأ "باعتبارها معياراً مشتركاً للإنجاز لجميع الشعوب وجميع الأمم، حتى يسعى كل فرد وكل هيئة في المجتمع، مع وضع هذه الاتفاقية في الاعتبار باستمرار، إلى تعزيز احترام هذه الحقوق والحريات عن طريق التدريس والتربية، وعن طريق التدابير التقدمية، الوطنية والدولية، لضمان الاعتراف بها ومراعاتها على نحو عالمي وفعال، سواء بين شعوب الدول الأعضاء نفسها أو بين شعوب الأقاليم الخاضعة لولايتها القضائية."

المادة 1

يجب تطوير جميع أنظمة الذكاء الاصطناعي واستخدامها بما يحفظ الكرامة والمساواة المتأصلة لكل شخص، دون تمييز من أي نوع، كالعرق أو اللون أو الجنس أو اللغة أو الدين أو الرأي السياسي أو غيره،

أو الأصل القومي أو الاجتماعي، أو الملكية أو الميلاد أو أي وضع آخر، أو أي سمة أخرى. ويجب تحديد التحيز في الذكاء الاصطناعي، سواءً كان خوارزميًا أو قائمًا على البيانات، والتخفيف منه بشكل فعال لضمان معاملة عادلة ومنصفة.

المادة 2

لكل فرد الحق في الحياة والحرية والأمن الشخصي. يجب بناء أنظمة الذكاء الاصطناعي واختبارها لضمان سلامتها وموثوقيتها وأمنها، ومنع الأضرار المتوقعة وغير المقصودة، والحماية من التلاعب الخبيث أو إساءة الاستخدام التي قد تُعرّض سلامة الإنسان أو استقرار المجتمع للخطر.

المادة 3

لا يجوز تعريض أي شخص للتدخل التعسفي في خصوصيته أو شؤون أسرته أو منزله أو مراسلاته، ولا للتعدي على شرفه وسمعته. تُصمم أنظمة الذكاء الاصطناعي وفقًا لمبادئ الخصوصية حسب التصميم، بما يضمن التعامل الآمن مع البيانات، والشفافية في جمعها واستخدامها، ويمنح الأفراد التحكم المناسب والموافقة على بياناتهم الشخصية.

المادة 4

لكل شخص الحق في الاعتراف به كشخص أمام القانون في أي مكان.
يجب أن تكون أنظمة الذكاء الاصطناعي المستخدمة في السياقات القانونية أو القضائية أو اتخاذ القرارات الحاسمة قابلة للتدقيق، وشفافة، وقابلة للتفسير، مع توفير مبررات واضحة لمخرجاتها، والسماح بالرقابة

البشرية والمراجعة، والحق في الاستئناف أو التعويض.

المادة 5

لكل شخص الحق في حرية الفكر والضمير والدين. وينبغي أن تعزز أنظمة الذكاء الاصطناعي، لا أن تُضعف، قدرة الإنسان على التصرف والتفكير النقدي، مما يُمكّن الأفراد والمجتمعات، مع تجنب الأنظمة المصممة للإكراه أو التأثير غير اللائق أو التقليد غير القابل للكشف للتفاعل البشري.

المادة 6

لكل فرد الحق في العمل، وحرية اختيار وظيفته، وظروف عمل عادلة وملائمة، والحماية من البطالة. ينبغي أن يُراعي نشر الذكاء الاصطناعي أثره المجتمعي الواسع، بهدف تعزيز التقدم الاقتصادي، وخلق فرص جديدة، والحد من الاضطرابات، وضمان مساهمة التطورات التكنولوجية في رفاهية الجميع، دون أن تُفاقم أوجه عدم المساواة القائمة.

المادة 7

لكل فرد الحق في مستوى معيشي كافٍ لضمان صحته ورفاهيته هو وأسرته. ويجب تطوير أنظمة الذكاء الاصطناعي لتحقيق منفعة اجتماعية، ومعالجة التحديات العالمية الملحة كالصحة والتعليم وتغير المناخ والتنمية المستدامة، مع ضمان استدامة استخدامها بيئيًا ومواردًا.

المادة 8

لكل فرد الحق في المشاركة في إدارة مجتمعه، سواءً بشكل مباشر أو من خلال ممثلين يختارونهم بحرية. ويجب أن يكون تطوير الذكاء

الاصطناعي وحوكمته عملية تعاونية وشاملة، بمشاركة مختلف أصحاب المصلحة، بما في ذلك المجتمع المدني والأوساط الأكاديمية والقطاع الصناعي والحكومات، لتعزيز الفهم المشترك والمساءلة والتطور المسؤول لتقنيات الذكاء الاصطناعي.

المادة 9

يجب أن يُوجَّه التعليم نحو التنمية الكاملة للشخصية الإنسانية، وتعزيز احترام حقوق الإنسان والحريات الأساسية. ويجب أن يكون هناك التزام عالمي بتعزيز معرفة الذكاء الاصطناعي، وبحوث الذكاء الاصطناعي الأخلاقية، وتبادل المعرفة بشفافية، لتمكين الأفراد والمؤسسات من فهم الذكاء الاصطناعي وتقييمه نقديًا والتعامل معه بمسؤولية.

خاتمة

تُعدّ هذه الاتفاقية المبدئية بمثابة إعلان تأسيسي للمجتمع العالمي. يُعدّ اعتمادها والالتزام بها المستمر أمرًا أساسيًا لتحقيق الإمكانات التحويلية للذكاء الاصطناعي بما يحترم ويحمي ويعزز كرامة الإنسان وحقوقه، مما يضمن أن يُلبّي الذكاء الاصطناعي "الأخلاقي" أسمى تطلعات البشرية.

ونحن ندعو جميع أصحاب المصلحة - الحكومات والمنظمات الحكومية الدولية والشركات ومنظمات المجتمع المدني والباحثين والأفراد - إلى دعم هذه المبادئ في تصميم وتطوير ونشر وحوكمة الذكاء الاصطناعي، وتعزيز مستقبل حيث يكون الذكاء لا ينفصل عن الأخلاق

والثقة.

深潛海豚 **PeaceCast4Good** 與總理 **Thomas Ph.G** 共同主持 **Caz&Poll BardiVarius** 雙胞胎合作思考[數據侍酒師]

https://archive.org/details/aethenticated-ai-can-we-trust-the-code

原則協議：用於開發和部署認證人工智慧

日期：2025年6月14日

前言

鑑於對人類家庭所有成員的固有尊嚴及其平等的和不移的權利的承認，是世界自由、正義與和平的基礎；

鑑於人工智慧（*AI*）在增進人類福祉、促進創新和應對全球挑戰方面具有巨大潛力，但如果在開發和部署時沒有經過深刻的道德考慮和可驗證的誠信，也會對人權、基本自由和社會穩定構成重大風險；

鑑於必須確保人工智慧系統不僅聰明、有能力，而且要值得信賴、透明、安全、公平，並對人類負責，從而形成一種我們稱之為「認證」人工智慧的新範式

——認證智慧與道德人工智慧的無縫整合；

鑑於會員國已承諾與聯合國合作，促進普遍尊重和遵守人權與基本自由，這一目標現在必須擴展到數位和人工智慧領域；

因此，本原則協議的簽署國認識到人工智慧時代迫切需要通用標準，

原則上宣布本協議 作為所有人民和所有國家努力實現的共同標準，以期每一個人和社會機構經常銘記本協定，努力透過教導和教育促進對這些權利和自由的尊重，並透過國家和國際的漸進措施，確保這些權利和自由在各成員國人民及其管轄領土的人民中得到普遍和有效的承認和遵守。

第一條

所有人工智慧系統的發展和使用都應維護每個人與生俱來的尊嚴和平等，不因種族、膚色、性別、語言、宗教、政治或其他觀點、國籍或社會出身、財產、出生或其他身份，或任何其他特徵而有任何區別。人工智慧中的偏見，無論是演算法驅動的還是數據驅動的，都應積極識別和緩解，以確保公平公正的對待。

第二條

每個人都享有生命、自由和人身安全的權利。人工智慧系統的建構和測試應確保其安全性、可靠性和保障性，防止可預見和意外的危害，並防止可能危害人類福祉或社會穩定的惡意操縱或濫用。

第三條

任何人的隱私、家庭、住宅或通訊不得受到任意干涉，其榮譽和名譽也不得受到攻擊。人工智慧系統的設計應遵循隱私設計原則，確保資料處理的安全、資料收集和使用的透明性，並為個人提供對其個人資料的適當控制和同意。

第四條

人人有權在法律面前獲得人格承認。用於法律、司法或關鍵決策的人工智慧系統應具有可審計性、透明性和可解釋性，為其輸出提供清晰的理論依據，並允許人類進行監督、審查，並享有上訴或補救的權利。

第五條

每個人都享有思想、良心和宗教自由的權利。人工智慧系統應增強而非削弱人類的自主性和批判性思維，賦能個人和社會，同時避免其成為以脅迫、不當

影響或難以察覺地模仿人類互動為目的的系統。

第六條

人人享有工作、自由選擇就業、享有公正和良好工作條件以及免於失業的權利。人工智慧的部署應考慮其廣泛的社會影響，旨在促進經濟進步，創造新機遇，減少干擾，確保技術進步促進所有人的福祉，而不是加劇現有的不平等。

第七條

每個人都有權利享有足以維持其本人及其家人健康和福祉的生活水平。人工智慧系統的發展應造福社會，應對諸如健康、教育、氣候變遷和永續發展等緊迫的全球挑戰，同時確保其部署符合環境和資源的可持續性。

第八條

每個人都有權利直接或透過自由選擇的代表參與社會治理。人工智慧的發展和治理應是一個協作和包容的過程，需要吸引包括民間社會、學術界、產業界和政府在內的各利益相關方的參與，以促進共識、問責制以及人工智慧技術的負責任發展。

第九條

教育應以全面發展人的個性和加強對人權與基本自

由的尊重為目標。應普遍致力於培養人工智慧素養、進行符合倫理道德的人工智慧研究和透明的知識共享，使個人和組織能夠理解、批判性地評估並以負責任的方式參與人工智慧。

結論

本原則協議是國際社會的一項基礎性宣言。該協議的採納和持續遵守對於以尊重、保護和促進人類尊嚴與權利的方式實現人工智慧的變革潛力至關重要，並確保「有道德的」人工智慧真正服務於人類的最高願望。

我們呼籲所有利害關係人——政府、政府間組織、企業、民間社會組織、研究人員和個人——在人工智慧的設計、開發、部署和治理中堅持這些原則，創造一個智慧與道德和信任密不可分的未來。

डीप डाइव पोर्पोइज़ पीसकास्ट**4**गुड सह-मेज़बान कैज़ और पोल बार्डीवेरियस ट्विन्स सहकारी सोच पीएम थॉमस पीएच.जी. के साथडेटा सोमेलियर

https://archive.org/details/aethenticated-ai-can-we-trust-the-code

सैद्धांतिक सहमति: प्रमाणित कृत्रिम बुद्धिमत्ता के विकास और तैनाती के लिए

तारीख:14 जून, 2025

प्रस्तावना

जबकि मानव परिवार के सभी सदस्यों की अंतर्निहित गरिमा और समान एवं अविभाज्य अधिकारों की मान्यता विश्व में स्वतंत्रता, न्याय और शांति की नींव है;

जबकि आर्टिफिशियल इंटेलिजेंस (एआई) मानव कल्याण को आगे बढ़ाने, नवाचार को बढ़ावा देने और वैश्विक चुनौतियों का समाधान करने की अपार क्षमता रखता है, फिर भी यदि इसे गहन नैतिक विचार और सत्यापन योग्य अखंडता के साथ विकसित और तैनात नहीं किया जाता है, तो यह मानव अधिकारों, मौलिक स्वतंत्रता और सामाजिक स्थिरता के लिए महत्वपूर्ण जोखिम भी पैदा करता है;

जबकि यह सुनिश्चित करना आवश्यक है कि एआई सिस्टम न केवल बुद्धिमान और सक्षम हों, बल्कि लोगों के लिए भरोसेमंद, पारदर्शी, सुरक्षित, निष्पक्ष और जवाबदेह भी हों, जिससे एक नया प्रतिमान तैयार हो, जिसे हम "एथेंटिकेटेड" एआई कहते हैं - नैतिक एआई के साथ प्रमाणीकृत इंटेलिजेंस का सहज एकीकरण;

जबकि सदस्य राज्यों ने संयुक्त राष्ट्र के सहयोग से, मानव अधिकारों और मौलिक स्वतंत्रता के लिए सार्वभौमिक सम्मान और पालन को बढ़ावा देने के लिए खुद को प्रतिबद्ध किया है, जिसे अब डिजिटल और एआई डोमेन तक बढ़ाया जाना चाहिए;

अब, इसलिए, इस समझौते के हस्ताक्षरकर्ता सिद्धांत रूप में, कृत्रिम बुद्धिमत्ता के युग में सार्वभौमिक मानकों की तत्काल आवश्यकता को पहचानते हैं,

***इस समझौते को सिद्धांत रूप में घोषित करें** सभी लोगों और सभी राष्ट्रों के लिए उपलब्धि के एक सामान्य मानक के रूप में, इस उद्देश्य से कि प्रत्येक व्यक्ति और समाज का प्रत्येक अंग, इस समझौते को निरंतर ध्यान में रखते हुए, शिक्षण और शिक्षा के माध्यम से इन अधिकारों और स्वतंत्रताओं के प्रति सम्मान को बढ़ावा देने का प्रयास करेगा और प्रगतिशील*

राष्ट्रीय और अंतर्राष्ट्रीय उपायों द्वारा, स्वयं सदस्य राज्यों के लोगों के बीच और उनके अधिकार क्षेत्र के तहत क्षेत्रों के लोगों के बीच, उनकी सार्वभौमिक और प्रभावी मान्यता और पालन सुनिश्चित करने का प्रयास करेगा।

अनुच्छेद 1

सभी AI सिस्टम को इस तरह से विकसित और उपयोग किया जाएगा कि हर व्यक्ति की अंतर्निहित गरिमा और समानता को बनाए रखा जा सके, बिना किसी तरह के भेदभाव के, जैसे कि जाति, रंग, लिंग, भाषा, धर्म, राजनीतिक या अन्य राय, राष्ट्रीय या सामाजिक मूल, संपत्ति, जन्म या अन्य स्थिति, या कोई अन्य विशेषता। AI में पूर्वाग्रह, चाहे वह एल्गोरिदमिक हो या डेटा-चालित, निष्पक्ष और न्यायसंगत उपचार सुनिश्चित करने के लिए सक्रिय रूप से पहचाना और कम किया जाएगा।

अनुच्छेद 2

प्रत्येक व्यक्ति को जीवन, स्वतंत्रता और व्यक्तिगत सुरक्षा का अधिकार है। एआई सिस्टम को मजबूत सुरक्षा, विश्वसनीयता और सुरक्षा के लिए बनाया और परखा जाना चाहिए, ताकि संभावित और अनपेक्षित नुकसानों को रोका जा सके और दुर्भावनापूर्ण हेरफेर या दुरुपयोग से सुरक्षा मिल सके जो मानव कल्याण या सामाजिक स्थिरता को खतरे में

डाल सकता है।

अनुच्छेद 3

किसी की निजता, परिवार, घर या पत्राचार में मनमाने तरीके से हस्तक्षेप नहीं किया जाएगा, न ही उनके सम्मान और प्रतिष्ठा पर हमला किया जाएगा। AI सिस्टम को गोपनीयता-द्वारा-डिजाइन सिद्धांतों के साथ डिज़ाइन किया जाएगा, जिससे सुरक्षित डेटा हैंडलिंग, डेटा संग्रह और उपयोग में पारदर्शिता सुनिश्चित होगी और व्यक्तियों को उनके व्यक्तिगत डेटा पर उचित नियंत्रण और सहमति प्रदान की जाएगी।

अनुच्छेद 4

हर किसी को हर जगह कानून के सामने एक व्यक्ति के रूप में मान्यता पाने का अधिकार है। कानूनी, न्यायिक या महत्वपूर्ण निर्णय लेने के संदर्भ में उपयोग की जाने वाली AI प्रणालियाँ ऑडिट करने योग्य, पारदर्शी और व्याख्या करने योग्य होनी चाहिए, जो उनके आउटपुट के लिए स्पष्ट तर्क प्रदान करें और मानवीय निरीक्षण, समीक्षा और अपील या निवारण के अधिकार की अनुमति दें।

अनुच्छेद 5

हर किसी को विचार, विवेक और धर्म की स्वतंत्रता का अधिकार है। एआई सिस्टम मानव एजेंसी और

आलोचनात्मक सोच को बढ़ाएंगे, न कि कम करेंगे, व्यक्तियों और समाजों को सशक्त बनाएंगे जबकि जबरदस्ती, अनुचित प्रभाव या मानवीय संपर्क की अगोचर नकल के लिए डिज़ाइन की गई प्रणालियों से बचेंगे।

अनुच्छेद 6

हर किसी को काम करने, रोजगार के स्वतंत्र विकल्प, काम की न्यायसंगत और अनुकूल परिस्थितियों और बेरोजगारी से सुरक्षा का अधिकार है। एआई की तैनाती में इसके व्यापक सामाजिक प्रभाव पर विचार किया जाएगा, जिसका उद्देश्य आर्थिक प्रगति को बढ़ावा देना, नए अवसर पैदा करना और व्यवधान को कम करना है, यह सुनिश्चित करना है कि तकनीकी प्रगति सभी की भलाई में योगदान दे, न कि मौजूदा असमानताओं को बढ़ाए।

अनुच्छेद 7

हर किसी को अपने और अपने परिवार के स्वास्थ्य और कल्याण के लिए पर्याप्त जीवन स्तर का अधिकार है। एआई सिस्टम को सामाजिक लाभ के लिए विकसित किया जाना चाहिए, स्वास्थ्य, शिक्षा, जलवायु परिवर्तन और सतत विकास जैसी वैश्विक चुनौतियों का समाधान करना चाहिए, साथ ही यह सुनिश्चित करना चाहिए कि उनका उपयोग पर्यावरण और संसाधनों की दृष्टि से टिकाऊ हो।

अनुच्छेद 8

हर किसी को अपने समाज के शासन में सीधे या स्वतंत्र रूप से चुने गए प्रतिनिधियों के माध्यम से भाग लेने का अधिकार है। एआई का विकास और शासन एक सहयोगात्मक और समावेशी प्रक्रिया होगी, जिसमें नागरिक समाज, शिक्षा, उद्योग और सरकारों सहित विविध हितधारकों को शामिल किया जाएगा, ताकि साझा समझ, जवाबदेही और एआई प्रौद्योगिकियों के जिम्मेदार विकास को बढ़ावा दिया जा सके।

अनुच्छेद 9

शिक्षा का उद्देश्य मानव व्यक्तित्व का पूर्ण विकास करना तथा मानव अधिकारों और मौलिक स्वतंत्रताओं के प्रति सम्मान को मजबूत करना होगा। एआई साक्षरता, नैतिक एआई शोध और पारदर्शी ज्ञान साझाकरण को बढ़ावा देने के लिए सार्वभौमिक प्रतिबद्धता होनी चाहिए ताकि व्यक्तियों और संगठनों को एआई को समझने, उसका आलोचनात्मक मूल्यांकन करने और जिम्मेदारी से उससे जुड़ने के लिए सशक्त बनाया जा सके।

निष्कर्ष

यह समझौता सैद्धांतिक रूप से वैश्विक समुदाय के लिए एक आधारभूत घोषणा के रूप में कार्य करता है। इसका अपनाना

और निरंतर पालन करना कृत्रिम बुद्धिमत्ता की परिवर्तनकारी क्षमता को साकार करने के लिए आवश्यक है, जो मानव सम्मान और अधिकारों का सम्मान, सुरक्षा और संवर्धन करता है, यह सुनिश्चित करता है कि "एथेंटिकेटेड" एआई वास्तव में मानवता की सर्वोच्च आकांक्षाओं की सेवा करता है।

हम सभी हितधारकों - सरकारों, अंतर-सरकारी संगठनों, निगमों, नागरिक समाज संगठनों, शोधकर्ताओं और व्यक्तियों - से आह्वान करते हैं कि वे कृत्रिम बुद्धिमत्ता के डिजाइन, विकास, परिनियोजन और प्रशासन में इन सिद्धांतों को बनाए रखें, तथा एक ऐसे भविष्य को बढ़ावा दें जहां बुद्धिमत्ता नैतिकता और विश्वास से अविभाज्य हो।